

INVARIANT TORI IN HAMILTONIAN SYSTEMS WITH HIGH ORDER PROPER DEGENERACY

YUECAI HAN, YONG LI, AND YINGFEI YI

ABSTRACT. We study the existence of quasi-periodic, invariant tori in a nearly integrable Hamiltonian system of high order proper degeneracy, i.e., the integrable part of the Hamiltonian involves several time scales and at each time scale the corresponding Hamiltonian depends on only part of the action variables. Such a Hamiltonian system arises frequently in problems of celestial mechanics, for instance, in perturbed Kepler problems like the restricted and non-restricted 3-body problems and spatial lunar problems in which several bodies with very small masses are coupled with two massive bodies and the nearly integrable Hamiltonian systems naturally involve different time scales. Using KAM method, we will show under certain higher order non-degenerate conditions of Bruno-Rüssmann type that the majority of quasi-periodic, invariant tori associated with the integrable part will persist after the non-integrable perturbation. This actually concludes the KAM metric stability for such a properly degenerate Hamiltonian system.

1. INTRODUCTION

In this paper, we study the existence of quasi-periodic motions in a nearly integrable, properly degenerate Hamiltonian system where the proper degeneracy occurs in an arbitrarily high order. To be more precise, associated with action-angle variables $(y, x) \in R^n \times T^n$ and the standard symplectic structure $dy \wedge dx$, we consider, for each parameter $\varepsilon > 0$ sufficiently small, a real analytic Hamiltonian of the form

$$(1.1) \quad \mathcal{H}(x, y, \varepsilon) = h_0(y^{n_0}) + \varepsilon^{m_1} h_1(y^{n_1}) + \cdots + \varepsilon^{m_a} h_a(y^{n_a}) + \varepsilon^{m_a+1} p(x, y, \varepsilon),$$

where, for given positive integers a, m , and l , $n_i, i = 0, 1, \dots, a, m_j, j = 1, 2, \dots, a$, are positive integers satisfying $n_0 \leq n_1 \leq \cdots \leq n_a =: n, m_1 \leq m_2 \leq \cdots \leq m_a =: m, y^{n_i} = (y_1, \dots, y_{n_i})^\top$, for $i = 1, 2, \dots, a$, and P depends on ε smoothly. The Hamiltonian (1.1) is clearly nearly integrable when the parameter ε is sufficiently small.

Hamiltonians of the form (1.1) arise frequently in problems of celestial mechanics, for instance, in perturbed Kepler problems like the restricted and non-restricted 3-body problems and spatial lunar problems in which several bodies with very small masses are coupled with two massive bodies and the nearly integrable Hamiltonian systems naturally involve different time scales (see [28] and references therein). After certain regularization and normalization (see e.g., [9, 14, 21]), the Hamiltonians typically have the form (1.1), for which proper degeneracies, mainly due to the super-integrability of the Keplerian, usually occur in a way that for each time scale of order ε^{m_j} , the normalized Hamiltonian h_j is only a function of the first n_j action variables for some positive integers m_j, n_j . Hence they are properly degenerate in the sense of Arnold ([1, 2]).

2000 *Mathematics Subject Classification.* Primary 37J40, 70H08.

Key words and phrases. Invariant tori, high order proper degeneracy, KAM theory.

The first author is partially supported by NSFC Grant 10601019 and Chinese Postdoctoral Science Foundation. The second author is partially supported by NSFC Grant 10531050, National 973 project of China 2006CD805903, SRFDP Grant 20040183030, and 985 Program of Jilin University. The third author is partially supported by NSF grant DMS0708331, NSFC Grant 10428101, and a Changjiang Scholarship from Jilin University.

The existence of quasi-periodic motions for properly degenerate Hamiltonian (1.1) was first shown by Arnold ([1]) for the case $a = m = 1$ under a so-called degeneracy-removing condition that $h_0 + \varepsilon h_1$ satisfies either the Kolmogorov or iso-energetic non-degenerate condition. Such a degeneracy-removing condition is known to satisfy in many planar or restricted 3-body problems and n -body problems, leading to the existence of quasi-periodic invariant tori (see [2, 3, 4, 6, 10, 11, 14, 15, 19, 20] and references therein). However, it is also known that in many cases of perturbed Kepler problems, the leading order of the perturbed Hamiltonian is insufficient to remove the degeneracy. For instance, the normalized Hamiltonian associated with the spatial lunar problem considered by Sommer ([28]) has the form

$$(1.2) \quad H = J_1 + \varepsilon h_1(J_1, J_2) + \varepsilon^2 h_2(J_1, J_2, J_3, \mu) + \varepsilon^l P(J, \phi, \varepsilon), \quad J = (J_1, J_2, J_3) \in R^3, \phi \in T^3,$$

where $l > 3$ is a real number. This Hamiltonian actually involves three time scales (i.e. $a = m = 2$ in (1.1)). As shown in [28], besides Arnold's singularity-removing condition imposed on the $O(\varepsilon)$ order term h_1 , the existence of quasi-periodic invariant tori for (1.2) requires a further singularity-removing condition of Kolmogorov type imposing on the $O(\varepsilon^2)$ order term h_2 .

Motivated by applications arising in a broader class of perturbed Kepler problems, the goal of this work is to present a KAM type of result for Hamiltonians of type (1.1) by taking into account of higher order singularity-removing conditions.

To set up the problem, we consider the Hamiltonian (1.1) in a bounded closed region $G \times T^n \subset R^n \times T^n$. It is clear that for each ε the integrable part of (1.1):

$$(1.3) \quad N_\varepsilon(y) = h_0(y^{n_0}) + \varepsilon^{m_1} h_1(y^{n_1}) + \cdots + \varepsilon^{m_a} h_a(y^{n_a}),$$

admits a family of invariant n -tori $T_\xi^\varepsilon = \{\xi\} \times T^n$, with linear flows $\{x_0 + \omega^\varepsilon(\xi)t\}$, $\xi \in G$, where for each $\xi \in G$,

$$\omega^\varepsilon(\xi) = \nabla N_\varepsilon(\xi)$$

is the frequency vector of the n -torus T_ξ^ε . When $\omega^\varepsilon(\xi)$ is non-resonant, the n -torus T_ξ^ε becomes quasi-periodic with slow and fast frequencies of different scales. Adopting the terminology of Arnold ([1, 2]), we refer the integrable part N_ε and its associated tori $\{T_\xi^\varepsilon\}$ to as the *intermediate Hamiltonian* and *intermediate tori* respectively.

Let $\hat{y}^{n_i} = (y_{n_{i-1}+1}, \dots, y_{n_i})^\top$, $i = 0, 1, \dots, a$, where $n_{-1} = 0$ (hence $\hat{y}^{n_0} = y^{n_0}$), and define

$$(1.4) \quad \Omega = (\nabla_{\hat{y}^{n_0}} h_0(y^{n_0}), \dots, \nabla_{\hat{y}^{n_a}} h_a(y^{n_a})),$$

where for each $i = 0, 1, \dots, a$, $\nabla_{\hat{y}^{n_i}}$ denotes the gradient with respect to \hat{y}^{n_i} . We assume the following high order, degeneracy-removing condition of Bruno-Rüssmann type:

A) *There is a positive integer N such that*

$$\text{Rank}\{\partial_y^\alpha \Omega(y); \quad 0 \leq |\alpha| \leq N\} = n, \quad \forall y \in \Omega.$$

We note that the condition A) above is equivalent to the following condition:

A') *For each $i = 0, 1, \dots, a$, there is a positive integer N_i such that*

$$\text{Rank}\left\{\frac{\partial^\alpha h_i(y^{n_i})}{\partial(\hat{y}^{n_i})^\alpha}; \quad 1 \leq |\alpha| \leq N_i\right\} = n_i - n_{i-1}, \quad \forall y \in \Omega.$$

We will prove the following

Theorem (Main Result). *Assume the condition A) and let $0 < \delta < \frac{1}{5}$ be given. Then there exists an $\varepsilon_0 > 0$ and a family of Cantor sets $G_\varepsilon \subset G$, $0 < \varepsilon < \varepsilon_0$, with $|G \setminus G_\varepsilon| = O(\varepsilon^{\frac{1}{N}})$, such that each $\xi \in G_\varepsilon$ corresponds to a real analytic, invariant, quasi-periodic n -torus \hat{T}_ξ^ε of the Hamiltonian (1.1) which is slightly deformed from the intermediate n -torus T_ξ^ε . Moreover, the family $\{\hat{T}_\xi^\varepsilon : \xi \in G_\varepsilon, 0 < \varepsilon < \varepsilon_0\}$ varies Whitney smoothly.*

Remark. 1) *Using arguments in [8], the above theorem also holds on a submanifold M of R^n if the condition A) is only assumed for $\xi \in M$ (e.g., M is a fixed energy surface). This in particular leads to an iso-energetic version of the theorem (see [8] for detail). One can further show the partial preservation of frequency components for the perturbed tori in the above theorem. More precisely, let i_1, \dots, i_{n_*} be the row indexes of a non-singular principal minor of the matrix $\partial\Omega$ on G . Then the i_1, \dots, i_{n_*} components of each perturbed toral frequency remain the same as the corresponding ones of the associated unperturbed toral frequency.*

2) *Differing from the case for a usual nearly integrable Hamiltonian system, the excluding measure for the existence of quasi-periodic invariant tori in the properly degenerate case is of a fairly large order of $\varepsilon^{\frac{1}{N}}$ for a pre-fixed small positive constant δ , as shown in the theorem above. This is mainly caused by a normal form reduction which pushes the perturbation to an order higher than $\varepsilon^{Nb+\delta}$ for*

$$(1.5) \quad b = \sum_{i=1}^a m_i(n_i - n_{i-1}),$$

for which the domain G needs to shrink by an order of $\varepsilon^{\frac{1}{N}}$ in measure. This is necessary for general properly degenerate Hamiltonian systems like (1.1) in order for the standard KAM iterations to apply (see the discussion below). However, if the perturbation in (1.1) is already in an order of $O(\varepsilon^{Nb+\delta})$, then a normal form reduction will not be necessary, and the excluding measure for the existence of quasi-periodic invariant tori can be improved to an order of ε^b (see the measure estimate in Section 3). Indeed, this is the case for (1.2) because $l > b = 3$ and $N = 1$ there. We note that in the case $l = 3$ in (1.2), direct KAM iterations are not applicable. Instead, one can apply the theorem above to obtain a nearly full measure set of KAM tori with the excluding measure in an order of ε^δ for some pre-fixed small positive constant δ .

3) *In applications, verification of the condition A) should rely on certain a priori regularization or normalization procedures which add higher order averaged terms to the properly degenerate part until the degeneracy-removing condition A) is satisfied. Such an averaging procedure can be made general if lower dimensional tori are considered (see [13]) but it can be very delicate for the case of full dimensional tori (see [28] for a complete treatment of the spatial lunar problem).*

For a usual nearly integrable Hamiltonian system

$$H(x, y) = N(y) + \varepsilon P(x, y), \quad (y, x) \in G \times T^n \subset R^n \times T^n,$$

the majority existence of invariant, quasi-periodic n -tori is asserted by the classical KAM theorem under the Kolmogorov non-degenerate condition that $\partial\omega(y)$, where $\omega(y) = \nabla N(y)$, is non-singular over G . The same was shown to hold by Bruno ([5]) under the Bruno non-degenerate condition that

$$\text{Rank}(\omega(y), \partial\omega(y)) = n, \quad y \in G.$$

The weakest condition guaranteeing such persistence was given by Rüssmann ([25]) under the Rüssmann non-degenerate condition that $\omega(G)$ should not lie in any $n - 1$ dimensional subspace (see also [7] for a similar geometric condition). KAM type of theorems under the Rüssmann non-degenerate condition were shown in [26, 30]. In particular, it was shown in [30] (see also [29]) that

the Rüssmann non-degenerate condition is equivalent to the condition A) above with respect to the present frequency map ω . We refer the readers to [8, 12, 16, 17, 18, 24, 27] for more KAM type of results under Rüssmann non-degenerate conditions.

Unfortunately, these results as well as their proofs do not apply to the properly degenerate Hamiltonian system (1.1) directly, simply because the order of its non-integrable perturbation is not high enough for the usual KAM iterations to carry over. Due to the nature of the proper degeneracy in (1.1), it is not hard to see that a possible KAM iteration for the Hamiltonian would have to be carried out over a frequency domain containing Diophantine frequencies of type (γ, τ) , for $\tau > \max\{(N+1)N-1, (n+1)n-1\}$ and $\gamma \sim \varepsilon^{Nb}$, where b is as in (1.5). This would automatically require a perturbation order that is higher than ε^{Nb} .

To overcome this obstacle, a crucial idea in the proof of our main result is to first obtain a normal form for (1.1) by conducting finitely many steps of KAM iterations on relatively small domains so that the non-integrable perturbation is pushed into a sufficiently high order. We will do so in Section 2 by adopting a quasi-linear KAM iterative scheme introduced in [16] which involves solving a system of quasi-linear homological equations at each KAM step instead of linear ones. Our main result will be proved in Section 3 by performing a linear KAM scheme for infinite steps.

Throughout the paper, unless specified otherwise, we will use the same symbol $|\cdot|$ to denote an equivalent (finite dimensional) vector norm and its induced matrix norm, absolute value of functions, and measure of sets etc., and use $|\cdot|_D$ to denote the sup-norm of functions on a domain D . For any $\tilde{r}, \tilde{s} > 0$, we let

$$D(\tilde{r}, \tilde{s}) = \{(x, y) : |\operatorname{Im} x| < \tilde{r}, |y| < \tilde{s}\}$$

be the (\tilde{r}, \tilde{s}) -complex neighborhood of $T^n \times \{0\} \subset T^n \times \mathbb{R}^n$, and

$$D(\tilde{s}) = \{y : |y| < \tilde{s}\}$$

be the \tilde{s} -complex neighborhood of $\{0\} \subset \mathbb{R}^n$.

2. REDUCTION TO NORMAL FORM

As usual, the translations $y \rightarrow y + \xi, x \rightarrow x, \xi \in G =: G_0$, transform (1.1) into a smooth family of real analytic Hamiltonians

$$(2.1) \quad \begin{aligned} \mathcal{H}^0 &= \mathcal{N}^0(y, \xi, \varepsilon) + \varepsilon^{m_a} \mathcal{P}^0(x, y, \xi, \varepsilon), \\ \mathcal{N}^0 &= N_\varepsilon(y + \xi) = e^0(\xi, \varepsilon) + \langle \omega^0(\xi, \varepsilon), y \rangle + h^0(y, \xi, \varepsilon) \end{aligned}$$

parametrized by $\xi \in G_0$, where $\omega^0 = \omega_\varepsilon$, $h^0 = O(|y^{n_0}|^2 + \varepsilon^{m_1}|y^{n_1}|^2 + \cdots + \varepsilon^{m_a}|y^{n_a}|^2)$, and $\mathcal{P}^0 = \varepsilon p(x, y + \xi, \varepsilon)$. It is clear that ω^0 has the form

$$\omega^0 = (\omega_0^0, \varepsilon^{m_1} \omega_1^0, \dots, \varepsilon^{m_a} \omega_a^0),$$

where $\omega_i^0 = \nabla_{\hat{y}^{n_i}} h_i(\xi^{n_i}) + O(\varepsilon^{m_{i+1}-m_i})$ for $i = 0, 1, \dots, a-1$ and $\omega_a^0 = \nabla_{\hat{y}^{n_a}} h_a(\xi^{n_a})$. We denote

$$\Omega^0 =: \Omega = (\omega_0^0, \omega_1^0, \dots, \omega_a^0).$$

We will derive a desired normal form for the Hamiltonian (2.1) via finite steps of KAM iterations using the quasi-linear iterative scheme introduced in [16]. As to be seen later, the term ε^{m_a} in the perturbation plays an important role during the iterations in controlling derivatives of the transformations. Hence the Hamiltonian (2.1) cannot be rescaled to include the term h^0 into the perturbation, which requires that each KAM iteration keeps a similar term in the integrable part. This is indeed one of the advantages of the quasi-linear scheme.

For the remaining part of the paper, all derivatives with respect to the parameter ξ should be understood in the sense of Whitney.

For the fixed $0 < \delta < \frac{1}{5}$ prescribed in the main result, we let $\gamma_0 = \varepsilon^\delta$, $s_0 = \varepsilon^{2\delta}$, $\mu_0 = \varepsilon^{1-5\delta}$. Also let $0 < r_0 < 1$ be given such that the Hamiltonian (2.1) is real analytic in $D(r_0, s_0)$. Then it is easy to see that

$$|\partial_\xi^l \mathcal{P}^0|_{D(r_0, s_0) \times G_0} < \gamma_0 s_0^2 \mu_0, \quad |l| \leq N.$$

Our normal form theorem states as follows.

Normal Form Theorem. *Consider the Hamiltonian (2.1) under the condition A). Then as $\varepsilon > 0$ sufficiently small, there exist a subsets $G_* \subset G_0$, with $|G_0 \setminus G_*| = O(\varepsilon^{\frac{\delta}{N}})$, and a smooth family of canonical, real analytic transformations $\Phi^* : \mathcal{D}(r_0, s_0) \rightarrow \mathcal{D}(r_*, s_*)$, where $r_* = \frac{r_0}{2}$ and $s_* = O(\varepsilon^{\frac{1}{5} + \frac{6\delta}{5}})$, such that the transformed Hamiltonian reads*

$$(2.2) \quad \mathcal{H}^* = \mathcal{H}^0 \circ \Phi^* = e^*(\xi, \varepsilon) + \langle \omega^*(\xi, \varepsilon), y \rangle + h^*(y, \xi, \varepsilon) + \mathcal{P}^*(x, y, \xi, \varepsilon),$$

where $h^* = O(|y|^2)$, ω^* has the form $\omega^* = (\omega_0^*, \varepsilon^{m_1} \omega_1^*, \dots, \varepsilon^{m_a} \omega_a^*)$ with ω_i being an $n_i - n_{i-1}$ dimensional vector for each $i = 0, 1, \dots, a$ respectively, and \mathcal{P}^* satisfies

$$(2.3) \quad |\partial_\xi^l \mathcal{P}^*|_{D(r_*, s_*) \times G_*} \leq \varepsilon \gamma_*^{2(N+6)} s_* \mu_*^2, \quad |l| \leq N$$

with $\gamma_* = \varepsilon^b$, $\mu_* = \varepsilon^{\frac{m_a + \delta}{2}}$. Moreover, if we denote $\Omega^* = (\omega_0^*, \omega_1^*, \dots, \omega_a^*)$, then

$$(2.4) \quad |\partial_\xi^l \Omega^* - \partial_\xi^l \Omega^0|_{G_*} \leq \varepsilon^{1-\frac{\delta}{2}}, \quad |l| \leq N.$$

We will prove the Normal Form Theorem inductively via a finite sequence of quasi-linear iterations. Suppose that at a ν th-step, we have obtained the following smooth family of real analytic Hamiltonians

$$(2.5) \quad \begin{aligned} \mathcal{H} &= \mathcal{N} + \varepsilon^{m_a} \mathcal{P}, \\ \mathcal{N} &= e(\xi, \varepsilon) + \langle \omega(\xi, \varepsilon), y \rangle + h(y, \xi, \varepsilon), \end{aligned}$$

where $(x, y) \in D(r, s)$ for some $0 < r = r < r_0$, $0 < s = s < s_0$, $\xi \in G$ with $G \subset R^n$ being a bounded region, ω has the form

$$\omega(\xi, \varepsilon) = (\omega_0, \varepsilon^{m_1} \omega_1, \dots, \varepsilon^{m_a} \omega_a)$$

with ω_i 's being an $n_i - n_{i-1}$ dimensional vectors for each $i = 0, 1, \dots, a$ respectively, h has the form

$$h(y, \xi, \varepsilon) = O(|y^{n_0}|^2 + \varepsilon^{m_1} |y^{n_1}|^2 + \dots + \varepsilon^{m_a} |y^{n_a}|^2),$$

and

$$(2.6) \quad |\partial_\xi^l \mathcal{P}|_{D(r, s) \times G} \leq \gamma_0 s^2 \mu, \quad |l| \leq N,$$

for some $0 < \mu \leq \mu_0$.

For “+” =: $\nu + 1$, we will find a symplectic transformation Φ^+ , which, on a small phase domain $D(r_+, s_+)$ and a smaller parameter domain G_+ , transforms (2.5) into a family of Hamiltonians

$$\mathcal{H}^+ = \mathcal{H} \circ \Phi^+ = \mathcal{N}^+ + \varepsilon^{m_a} \mathcal{P}^+$$

in the $(\nu + 1)$ th-step which enjoy similar properties as (2.5) but with a much smaller non-integrable perturbation \mathcal{P}^+ .

All constants $c_1 - c_5$ below are independent of iteration process. For simplicity, we will use c to denote any intermediate positive constant which is independent of the iteration process.

Define

$$\Omega = (\omega_0, \omega_1, \dots, \omega_a)$$

and let

$$\begin{aligned} r_+ &= \frac{r}{2} + \frac{r_0}{4}, \\ s_+ &= \frac{1}{4}\alpha s, \quad \alpha = \mu^{\frac{1}{3}}, \\ K_+ &= ([\log \frac{1}{\mu}] + 1)^3, \\ D_{\frac{i}{4}\alpha} &= D(r_+ + \frac{i-1}{4}(r-r_+), \frac{i}{4}\alpha s), \quad i = 1, 2, 3, 4, \\ \hat{D}(s) &= D(r_+ + \frac{3}{4}(r-r_+), s), \\ G_+ &= \{\xi \in G : |\langle k, \Omega(\xi, \varepsilon) \rangle| > \frac{\gamma_0}{|k|^\tau}, \quad 0 < |k| \leq K_+\}. \end{aligned}$$

Hereafter, we let $\tau > \max\{(N+1)N-1, (n+1)n-1\}$ be fixed.

We consider the truncation

$$(2.7) \quad \mathcal{R} = \sum_{|k| \leq K_+, |j| \leq 2} p_{kj} y^j e^{\sqrt{-1}\langle k, x \rangle}$$

of the Taylor-Fourier series

$$\mathcal{P} = \sum_{k \in Z^n, j \in Z_+^n} p_{kj} y^j e^{\sqrt{-1}\langle k, x \rangle}.$$

Lemma 2.1. *Assume*

$$\mathbf{H1)} \quad e^{-K_+ \frac{r-r_+}{4}} = o(\mu(r-r_+)^n).$$

Then there is a constant c_1 such that

$$\begin{aligned} |\partial_\xi^l(\mathcal{P} - \mathcal{R})|_{D_{\frac{3}{4}\alpha} \times G} &\leq c_1 \gamma_0 s^2 \mu^2, \\ |\partial_\xi^l \mathcal{R}|_{D_{\frac{3}{4}\alpha} \times G} &\leq c_1 \gamma_0 s^2 \mu \end{aligned}$$

for all $|l| \leq N$.

Proof. Write

$$\mathcal{P} = \mathcal{R} + I + II,$$

where

$$\begin{aligned} I &= \sum_{|k| > K_+, j \in Z_+^n} p_{kj} y^j e^{\sqrt{-1}\langle k, x \rangle}, \\ II &= \sum_{|k| \leq K_+, |j| \geq 3} p_{kj} y^j e^{\sqrt{-1}\langle k, x \rangle}. \end{aligned}$$

The standard Cauchy estimate yields that

$$\begin{aligned} |\partial_\xi^l I|_{\hat{D}(s) \times G} &\leq \sum_{|k| > K_+} |\partial_\xi^l \mathcal{P}|_{D(r,s) \times G} e^{-|k| \frac{r-r_+}{4}} \leq \gamma_0 s^2 \mu \sum_{\kappa=K_+}^{\infty} \kappa^n e^{-\kappa \frac{r-r_+}{4}} \\ &\leq \gamma_0 s^2 \mu \int_{K_+}^{\infty} t^n e^{-t \frac{r-r_+}{4}} dt \leq \gamma_0 s^2 \mu \frac{4^n n!}{(r-r_+)^n} e^{-K_+ \frac{r-r_+}{4}} \leq \gamma_0 s^2 \mu^2. \end{aligned}$$

It follows that

$$|\partial_\xi^l(\mathcal{P} - I)|_{\dot{D}(s) \times G} \leq |\partial_\xi^l \mathcal{P}|_{D(r,s) \times G} + |\partial_\xi^l I|_{\dot{D}(s) \times G} \leq 2\gamma_0 s^2 \mu.$$

Let \int be the obvious anti-derivative of $\frac{\partial^3}{\partial y^3}$. We have by Cauchy estimate that

$$\begin{aligned} |\partial_\xi^l II|_{D_{\frac{3}{4}\alpha} \times G} &\leq \frac{c}{s^3} \left| \int |\partial_\xi^l(\mathcal{P} - I - \mathcal{R})|_{\dot{D}(s) \times G} dy \right|_{D_{\frac{3}{4}\alpha}} = \frac{c}{s^3} \left| \int |\partial_\xi^l(\mathcal{P} - I)|_{\dot{D}(s) \times G} dy \right|_{D_{\frac{3}{4}\alpha}} \\ &\leq \frac{c}{s^3} \left| \int \gamma_0 s^2 \mu dy \right|_{D_{\frac{3}{4}\alpha}} \leq c\gamma_0 s^2 \mu^2. \end{aligned}$$

Thus,

$$(2.8) \quad |\partial_\xi^l(\mathcal{P} - \mathcal{R})|_{D_{\frac{3}{4}\alpha} \times G} \leq c\gamma_0 s^2 \mu^2,$$

and,

$$|\partial_\xi^l \mathcal{R}|_{D_{\frac{3}{4}\alpha} \times G} \leq |\partial_\xi^l(\mathcal{P} - \mathcal{R})|_{D_{\frac{3}{4}\alpha} \times G} + |\partial_\xi^l \mathcal{P}|_{D(r,s) \times G} \leq c\gamma_0 s^2 \mu.$$

□

We wish to average out all coefficients of \mathcal{R} by constructing a symplectic transformation as the time-1 map ϕ_F^1 of the flow generated by a Hamiltonian F of the form

$$F = \sum_{0 < |k| \leq K_+, |j| \leq 2} f_{kj} y^j e^{\sqrt{-1}\langle k, x \rangle}.$$

Consider the homological equation

$$(2.9) \quad \{\mathcal{N}, F\} + \varepsilon^{m_a}(\mathcal{R} - [\mathcal{R}]) = 0,$$

where $[\mathcal{R}] = \int_{T^n} \mathcal{R}(x, \cdot) dx$ is the average of \mathcal{R} .

Substituting the Taylor-Fourier series of F and \mathcal{R} into (2.9) yields

$$\begin{aligned} &- \sum_{0 < |k| \leq K_+, |j| \leq 2} \sqrt{-1}\langle k, \omega + \partial_y h \rangle f_{kj} y^j e^{\sqrt{-1}\langle k, x \rangle} \\ &= -\varepsilon^{m_a} \sum_{0 < |k| \leq K_+, |j| \leq 2} p_{kj} y^j e^{\sqrt{-1}\langle k, x \rangle}. \end{aligned}$$

By equating the coefficients above, we then obtain the following quasi-linear equations:

$$(2.10) \quad \sqrt{-1}\langle k, \omega(\xi, \varepsilon) + \partial_y h \rangle f_{kj} = \varepsilon^{m_a} p_{kj}, \quad |j| \leq 2, \quad 0 < |k| \leq K_+.$$

We note that, in general, $\partial_y h \neq 0$ and solutions f_{kj} of (2.10) are necessarily functions of y . This is a main difference between the quasi-linear scheme and the usual linear ones.

Lemma 2.2. *The following holds.*

1) *If*

$$\mathbf{H2)} \quad \max\{s, \varepsilon\} K_+^{\tau+1} = o(\gamma_0),$$

then the quasi-linear equations (2.10) can be uniquely solved on $D(s) \times G_+$ to obtain a function F which is real analytic in x, y and smooth in ξ , and moreover, there is a constant c_2 such that for all $l, i, j \in Z_+^n$ with $|i| + |j| \leq N + 1 - |l|$, $|l| \leq N$,

$$|\partial_\xi^l \partial_x^i \partial_y^j F|_{\hat{D}(s) \times G_+} \leq \begin{cases} c_2 \frac{\gamma_0^{N+3-|l|-|j|} s^{2-|j|} \mu}{(r-r_+)^{\chi}}, & |j| \leq 2; \\ c_2 \frac{\gamma_0^{N+3-|l|-|j|} \mu}{(r-r_+)^{\chi}}, & 2 \leq |j| \leq N+3-|l|, \end{cases}$$

where

$$\chi = (N+1)([\tau] + 1) + N + n + 1.$$

2) If we further assume

$$\mathbf{H3)} \quad s\alpha^4 = o((r-r_+)^{\chi+1}),$$

then

$$\phi_F^t : D_{\frac{\alpha}{2}} \rightarrow D_\alpha$$

is well defined for all $0 \leq t \leq 1$, and moreover, there is a constant c_3 such that for all $l, i, j \in Z_+^n$ with $|i| + |j| \leq N - |l|$, $|l| \leq N$,

$$(2.11) \quad |\partial_\xi^l \partial_x^i \partial_y^j (\phi_F^t - id)|_{\hat{D}(s) \times G_+} \leq \begin{cases} c_3 \frac{\gamma_0^{N+2-|l|-|j|} s^{2-|j|} \mu}{(r-r_+)^{\chi}}, & |j| \leq 1; \\ c_3 \frac{\gamma_0^{N+2-|l|-|j|} \mu}{(r-r_+)^{\chi}}, & 1 \leq |j| \leq N+2-|l| \end{cases}$$

for all $0 \leq t \leq 1$.

Proof. Let $(y, \xi) \in D(s) \times G_+$, $0 < |k| \leq K_+$, and denote

$$L_k = \sqrt{-1} \langle k, \omega(\xi) + \partial_y h \rangle.$$

We write $k = (k_0, k_1, \dots, k_a)$, where $k_i \in Z^{n_i - n_{i-1}}$ for each $i = 0, 1, \dots, a$ respectively. Let k_j , for some $j = 0, 1, \dots, a$, be the first nonzero components of k with respect to the splitting above. Then

$$\langle L_k, k \rangle = \varepsilon^{m_j} \langle \omega_j + O(|y^{n_j}|), k_j \rangle + \dots + \varepsilon^{m_a} \langle \omega_a + O(|y^{n_a}|), k_a \rangle.$$

By H2) and the definition of G_+ , we have

$$(2.12) \quad |\langle L_k, k \rangle| > \frac{\varepsilon^{m_j} \gamma_0}{|k|^\tau} - (\varepsilon^{m_j} O(s) + O(\varepsilon^{m_{j+1}}) K_+) > \frac{\varepsilon^{m_j} \gamma_0}{2|k|^\tau}.$$

Hence

$$\begin{aligned} |\partial_{\xi, y} \langle L_k, k \rangle^{-1}| &\leq |\langle L_k, k \rangle|^{-2} |\langle \partial_{\xi, y} L_k, k \rangle| \\ &\leq c \frac{|k|^{2\tau}}{\varepsilon^{2m_j} \gamma_0^2} \varepsilon^{m_j} |k| = c \frac{|k|^{2\tau+1}}{\varepsilon^{m_j} \gamma_0^2} \leq c \frac{|k|^{2\tau+1}}{\varepsilon^{m_a} \gamma_0^2}. \end{aligned}$$

It follows from induction that, for any $l, j \in Z_+^n$,

$$(2.13) \quad |\partial_\xi^l \partial_y^j \langle L_k, k \rangle^{-1}| \leq c \frac{|k|^{(l+j+1)\tau+l}}{\varepsilon^{m_a} \gamma_0^{l+j+1}}.$$

Now, by (2.12), L_k is non-vanishing on G_+ , i.e., the quasi-linear equations (2.10) are uniquely solvable on G_+ to yield solutions

$$(2.14) \quad f_{kj} = f_{kj}(y, \xi) = L_k^{-1} p_{kj},$$

$(y, \xi) \in D(s) \times G_+$, $0 < |k| \leq K_+$, $|j| \leq 2$. Since it is clear that $\bar{f}_{kj}(\bar{y}, \xi) = f_{-kj}(y, \xi)$ for all $|j| \leq 2$, $0 < |k| \leq K_+$, F is real analytic.

We note by Cauchy estimate that

$$|\partial_\xi^l p_{kj}|_{G_+} \leq |\partial_\xi^l \partial_y^j \mathcal{P}|_{D(r,s) \times G_+} e^{-|k|r} \leq \gamma_0^{N+3} s^{N-|j|} \mu e^{-|k|r}, \quad |l| \leq N+1, \quad |j| \leq 2.$$

Let

$$a(|j|) = \begin{cases} 2 - |j|, & \text{if } |j| \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

We have by (2.13) and (2.14) that

$$\begin{aligned} |\partial_\xi^l \partial_x^i \partial_y^j F| &\leq c \sum_{|j| \leq 2, 0 < |k| \leq K_+} |k|^i |\partial_y^j (\partial_\xi^l f_{kj} y^j)| e^{|k|(r_+ + \frac{3}{4}(r-r_+))} \\ &\leq c \gamma_0^{N+3-|l|-|j|} s^{a(|j|)} \mu \sum_{0 < |k| \leq K_+} |k|^\chi e^{-|k|\frac{r-r_+}{4}} \\ (2.15) \quad &\leq c \gamma_0^{N+3-|l|-|j|} s^{a(|j|)} \mu \int_1^\infty t^\chi e^{-t\frac{r-r_+}{4}} dt \leq \frac{\gamma_0^{N+3-|l|-|j|} s^{a(|j|)} \mu}{(r-r_+)^{\chi}}. \end{aligned}$$

Now assume H3) also holds. We write $\phi_F^t = (\phi_1^t, \phi_2^t)^\top$, where

$$(2.16) \quad \phi_{F1}^t(x, y) = x + \int_0^t F_y \circ \phi_F^u du,$$

$$(2.17) \quad \phi_{F2}^t(x, y) = y + \int_0^t F_x \circ \phi_F^u du.$$

Then for any $(x, y) \in D_{\frac{\alpha}{2}}$ and $t \in [0, 1]$, we have by (2.15) and H2) that

$$\begin{aligned} |\phi_{F1}^t(x, y) - x| &\leq |F_y|_{\hat{D}(s)} \leq c_2 \frac{s\mu}{(r-r_+)^{\chi}} < \frac{(r-r_+)}{2}, \\ |\phi_{F2}^t(x, y) - y| &\leq |F_x|_{\hat{D}(s)} \leq c_2 \frac{s^2\mu}{(r-r_+)^{\chi}} < \frac{\alpha s}{2}, \end{aligned}$$

i.e., $\phi_F^t(x, y) \in D_\alpha$.

The proof for (2.11) simply follows from (2.15) - (2.17).

□

Let $\Phi^+ = \phi_F^1$. Then the above lemma implies that for each $\xi \in G_+$, $\Phi^+ : D(r_+, s_+) \rightarrow D_{\frac{\alpha}{2}} \subset D(r, s)$ is well defined, symplectic, and real analytic. Now it is easy to see that

$$\begin{aligned} \mathcal{H}^+ &\equiv \mathcal{H} \circ \Phi^+ = \mathcal{N}^+ + \varepsilon^{m_a} \mathcal{P}^+, \\ \mathcal{N}^+ &= \mathcal{N} + [\mathcal{R}] = e^+ + \langle \omega^+, y \rangle + h^+, \end{aligned}$$

where

$$\begin{aligned} e^+ &= e + \varepsilon^{m_a} p_{00}, \\ \omega^+ &= \omega + \varepsilon^{m_a} (\{p_{0j}\}_{|j|=1}) =: (\omega_0^+, \varepsilon^{m_1} \omega_1^+, \dots, \varepsilon^{m_a} \omega_a^+), \\ h^+ &= h + \varepsilon^{m_a} \sum_{|j|=2} p_{kj} y^j, \\ (2.18) \quad \mathcal{P}^+ &= \int_0^1 \{\mathcal{R}_t, F\} \circ \phi_F^t dt + (\mathcal{P} - \mathcal{R}) \circ \phi_F^1, \end{aligned}$$

with ω_i^+ being an $n_i - n_{i-1}$ dimensional vector for each $i = 0, 1, \dots, a$ respectively, and

$$\mathcal{R}_t = (1-t)[\mathcal{R}] + t\mathcal{R}.$$

It is clear that $h^+ = O(|y^{n_0}|^2 + \varepsilon^{m_1}|y^{n_1}|^2 + \dots + \varepsilon^{m_a}|y^{n_a}|^2)$.

Define

$$\Omega^+ = (\omega_0^+, \omega_1^+, \dots, \omega_a^+).$$

Lemma 2.3. *Assume H1)-H3). Then the following holds.*

1) *There is a constant $c_4 > 0$ such that*

$$|\partial_\xi^l \Omega^+ - \partial_\xi^l \Omega|_{G_+} \leq c_4 \gamma_0 s \mu, \quad |l| \leq N.$$

2) *There is a constant c_5 such that*

$$|\partial_\xi^l P^+| \leq c_5 \frac{\gamma_0 s^2 \mu^2}{(r - r_+)^{2\chi}}, \quad |l| \leq N.$$

Proof. The proof of 1) is straightforward.

The proof of 2) follows from Lemmas 2.1, 2.2, (2.18), and direct computations. \square

Let

$$\mu_+ = c_0 \mu^{1+\delta},$$

where $c_0 = 16c_5$. If we assume

$$\mathbf{H4)} \quad \alpha^{1-3\delta} = o((r - r_+)^{2\chi}),$$

then it is clear that

$$|\partial_\xi^l P^+|_{D(r_+, s_+) \times G_+} \leq \gamma_0 s_+^2 \mu_+, \quad |l| \leq N.$$

This completes the $(\nu + 1)$ th step of iteration.

Proof of the Normal Form Theorem. We have used the following iterative sequences

$$\begin{aligned} r_\nu &= r_0 \left(1 - \sum_{i=1}^{\nu} \frac{1}{2^{i+1}}\right), \\ s_\nu &= \frac{1}{4} \alpha_{\nu-1} s_{\nu-1}, \\ \alpha_\nu &= \mu_\nu^{\frac{1}{3}}, \\ \mu_\nu &= c_0 \mu_{\nu-1}^{1+\delta}, \\ K_\nu &= ([\log(\frac{1}{\mu_{\nu-1}})] + 1)^3, \\ G_\nu &= \{\xi \in G_{\nu-1} : |\langle k, \Omega_{\nu-1}(\xi) \rangle| > \frac{\gamma_0}{|k|^\tau}, \quad 0 < |k| \leq K_\nu\}, \quad \nu = 1, 2, \dots \end{aligned}$$

Since

$$\begin{aligned} r_\nu - r_{\nu+1} &= \frac{r_0}{2^{\nu+2}}, \\ \mu_\nu &= c_0 \frac{(1+\delta)^{\nu-1}}{\delta} \varepsilon^{(1-5\delta)(1+\delta)^\nu}, \end{aligned} \tag{2.19}$$

it is easy to see that hypotheses H1), H3), H4), and part of the hypothesis H2), i.e.,

$$sK_+^{\tau+1} = o(\gamma_0)$$

hold for all ν . But the other part of the hypothesis H2), i.e.,

$$(2.20) \quad \varepsilon K_+^{\tau+1} = o(\gamma_0)$$

will only hold if the number of iterations is finite. In fact, if we take

$$\nu_* = \left\lceil \frac{\log(2(N+6)b+1) - \log(1-5\delta)}{\log(1+\delta)} \right\rceil + 1,$$

then it is easy to see that (2.20) holds as $\varepsilon \ll 1$ for all $\nu = 1, 2, \dots, \nu_*$. Hence the quasi-linear scheme can be performed inductively to generate a sequence of Hamiltonians

$$\begin{aligned} \mathcal{H}^\nu &= \mathcal{H}^{\nu-1} \circ \Phi^\nu = \mathcal{N}^\nu + \varepsilon^{m_a} \mathcal{P}^\nu, \\ \mathcal{N}^\nu &= e^\nu + \langle \omega^\nu, y \rangle + h^\nu \end{aligned}$$

(2.21)

defined on $D(r_\nu, s_\nu) \times G_\nu$, for all $\nu = 1, 2, \dots, \nu_*$.

Let $r_* = r_{\nu_*}$, $s_* = s_{\nu_*}^2$, $\gamma_* = \varepsilon^b$, $\mu_* = \varepsilon^{\frac{m_a+\delta}{2}}$, $\mathcal{H}^* = \mathcal{H}^{\nu_*}$, $e^* = e^{\nu_*}$, $\omega^* = \omega^{\nu_*}$, $\Omega^* = \Omega^{\nu_*}$, $h^* = h^{\nu_*}$, $\mathcal{P}^* = \varepsilon^{m_a} \mathcal{P}^{\nu_*}$, $G_* = G_{\nu_*}$, $\Phi^* = \Phi^{\nu_*}$. By (2.19), we can choose $\varepsilon \ll 1$ independent of ν such that

$$(2.22) \quad \mu_\nu < \varepsilon^{(1-5\delta)(1+\delta)^\nu}, \quad \nu = 1, 2, \dots, \nu_*.$$

It follows that

$$\mu_{\nu_*} \leq \varepsilon^{2(N+6)b+1} =: \varepsilon \gamma_*^{2(N+6)},$$

and hence

$$|\partial_\xi^l \mathcal{P}^*|_{D(r_*, s_*) \times G_*} \leq \varepsilon^{m_a} \gamma_0 s_{\nu_*}^2 \mu_{\nu_*} \leq \varepsilon \gamma_*^{2(N+6)} s_* \mu_*^2,$$

i.e., (2.3) holds. By Lemma 2.3 1), (2.4) also holds.

Since by condition A),

$$(2.23) \quad \text{Rank}\{\partial_\xi^\alpha \Omega^0; 0 \leq |\alpha| \leq N\} = n, \quad \forall \xi \in G_0,$$

Lemma 2.3 1) also implies that

$$\text{Rank}\{\partial_\xi^\alpha \Omega^\nu; 0 \leq |\alpha| \leq N\} = n, \quad \forall \xi \in G_\nu, \quad \nu = 1, 2, \dots, \nu_*.$$

It follows from the standard measure estimates involving Rüssmann conditions (see [30]) that

$$|G^{\nu-1} \setminus G^\nu| = O(\gamma_0) = O(\varepsilon^{\frac{\delta}{N}}), \quad \nu = 1, 2, \dots, \nu_*$$

(see also the measure estimate in the next section). Hence

$$|G^0 \setminus G^*| = \sum_{\nu=1}^{\nu_*} |G^{\nu-1} \setminus G^\nu| = O(\varepsilon^{\frac{\delta}{N}}).$$

This completes the proof. \square

3. PROOF OF THE MAIN RESULT

The rescaling

$$y \rightarrow \varepsilon \gamma_*^{N+6} \mu_* y, \quad \mathcal{H}_{\nu_*} \rightarrow \frac{\mathcal{H}_{\nu_*}}{\varepsilon \gamma_*^{N+6} \mu_*}$$

to the normal form (2.2) yields the Hamiltonian

$$H_0 =: \frac{\mathcal{H}_{\nu_*}}{\varepsilon \gamma_*^{N+6} \mu_*} = e_0 + \langle \omega_0, y \rangle + P_0$$

defined on $D(r_0, s_0) \times \Lambda_0$, where $r_0 =: r_*$, $s_0 =: s_*$, $\Lambda_0 = \bar{G}_*$, $\omega_0 = \omega^*$, and

$$P_0 = \frac{h^*(y, \varepsilon) + \mathcal{P}^*}{\varepsilon \gamma_*^{N+6} \mu_*}.$$

It is clear that

$$|\partial_\xi^l P_0|_{D(r_0, s_0) \times \Lambda_0} \leq \gamma_0 s_0 \mu_0, \quad |l| \leq N,$$

where $\gamma_0 =: \gamma_*^{N+6}$, $\mu_0 =: \mu_*$.

We consider the following sequences

$$\begin{aligned} r_\nu &= r_0 \left(1 - \sum_{i=1}^{\nu} \frac{1}{2^{i+1}}\right), \\ s_\nu &= \frac{1}{8} \alpha_{\nu-1} s_{\nu-1}, \\ \alpha_\nu &= \mu_\nu^{\frac{1}{2}}, \\ \mu_\nu &= c_0 \mu_{\nu-1}^{\frac{5}{4}}, \\ \gamma_\nu &= \gamma_0 \left(1 - \sum_{i=1}^{\nu} \frac{1}{2^{i+1}}\right), \\ K_\nu &= \left(\left\lfloor \log\left(\frac{1}{\mu_{\nu-1}}\right) \right\rfloor + 1\right)^{3\eta}, \end{aligned}$$

$\nu = 1, 2, \dots$, where $\eta > \frac{\log 2}{\log 5 - \log 4}$ is fixed.

The following iteration lemma and convergence result are special cases of those contained in [8].

Lemma 3.1. *Let ε be sufficiently small. Then the following holds for all $\nu = 1, 2, \dots$.*

- 1) *There is a sequence of smooth families of symplectic, real analytic, near identity transformations*

$$\Phi_\nu^\xi : D(r_\nu, s_\nu) \rightarrow D(r_{\nu-1}, s_{\nu-1}); \quad \xi \in \Lambda_\nu$$

such that

$$\begin{aligned} H_\nu &= H_{\nu-1} \circ \Phi_\nu^\xi =: N_\nu + P_\nu, \\ N_\nu &= e_\nu + \langle \omega_\nu, y \rangle, \end{aligned}$$

where

$$\begin{aligned}
 \Lambda_\nu &= \{\xi \in \Lambda_{\nu-1} : |\langle k, \omega_{\nu-1} \rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau}, \ 0 < |k| \leq K_\nu\} \\
 &= \{\xi \in \Lambda_{\nu-1} : |\langle k, \omega_{\nu-1}(\xi) \rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau}, \ K_{\nu-1} < |k| \leq K_\nu\}, \\
 (3.1) \quad |\partial_\xi^l \omega_\nu - \partial_\xi^l \omega_0|_{\Lambda_\nu} &\leq \gamma_0 \mu_0, \quad |l| \leq N, \\
 |\partial_\xi^l P_\nu|_{D_\nu \times \Lambda_\nu} &\leq \gamma_\nu s_\nu \mu_\nu, \quad |l| \leq N.
 \end{aligned}$$

2) *The Whitney extensions of*

$$\Psi^\nu =: \Phi_1^\xi \circ \Phi_2^\xi \circ \cdots \circ \Phi_\nu^\xi$$

converge C^N uniformly to a smooth family of symplectic maps, say, Ψ^∞ , on $D(\frac{r_0}{2}, \frac{s_0}{2}) \times \Lambda_*$, where

$$\Lambda_* = \bigcap_{\nu \geq 0} \Lambda_\nu,$$

such that

$$H_\nu = H_0 \circ \Psi_{\nu-1} \rightarrow H_\infty =: H_0 \circ \Psi^\infty = e_\infty + \langle \omega_\infty, y \rangle + P_\infty$$

with $e_\infty = \lim_{\nu \rightarrow \infty} e_\nu$, $\omega_\infty = \lim_{\nu \rightarrow \infty} \omega_\nu$, $P_\infty = \lim_{\nu \rightarrow \infty} P_\nu$, and moreover,

$$\partial_y^j P_\infty|_{D(\frac{r_0}{2}, 0) \times \Lambda_*} = 0, \quad |j| \leq 2.$$

Hence for each $\xi \in \Lambda_*$, $T^n \times \{0\}$ is an analytic invariant torus of H_∞ with Diophantine frequency $\omega_\infty(\xi)$ of type (γ_*, τ) for $\gamma_* = \lim_{\nu \rightarrow \infty} \gamma_\nu$.

We now estimate the measure $|\Lambda_0 \setminus \Lambda_*|$. For each $k \in Z^n \setminus \{0\}$ and $\nu = 0, 1, \dots$, we consider the set

$$R_k^{\nu+1} = \{\xi \in \Lambda_\nu : |g_k^\nu(\xi)| \leq \frac{\gamma_\nu}{|k|^{\tau+1}}\},$$

where

$$g_k^\nu(\xi) = \langle \frac{k}{|k|}, \omega_\nu \rangle.$$

Then

$$\Lambda_0 \setminus \Lambda_* = \bigcup_{\nu=0}^{\infty} \bigcup_{K_\nu < |k| \leq K_{\nu+1}} R_k^{\nu+1}.$$

By (2.4) and (3.1), we have

$$\frac{\partial^N g_k^\nu}{\partial \xi^N} = \sum_{i=0}^a \langle \frac{k}{|k|}, \varepsilon^{m_i} (\frac{\partial^N \omega_i^0}{\partial \xi^N} + O(\varepsilon^{1-\frac{\delta}{2}})) \rangle,$$

which, together with (2.23) implies that there is a constant $c > 0$ such that

$$|\frac{\partial^N g_k^\nu}{\partial \xi^N}|_{\Lambda_\nu} \geq c\varepsilon^b.$$

It then follows from Lemma 2.1 in [30] that

$$|R_k^{\nu+1}| \leq c(\varepsilon) \left(\frac{\varepsilon^{(N+6)b}}{|k|^{\tau+1}} \right)^{\frac{1}{N}}, \quad k \in Z^n \setminus \{0\}, \quad \nu = 0, 1, \dots,$$

where

$$c(\varepsilon) = 2(2 + 3 + \cdots + \frac{1}{n-1} + \frac{1}{c\varepsilon^b}).$$

Hence

$$(3.2) \quad |\Lambda_0 \setminus \Lambda_*| \leq \sum_{\nu=0}^{\infty} \sum_{K_\nu < |k| \leq K_{\nu+1}} |R_k^{\nu+1}| \leq d(\varepsilon) \varepsilon^{\frac{6b}{N}} \sum_{\nu=0}^{\infty} \sum_{K_\nu < |k| \leq K_{\nu+1}} \frac{1}{|k|^{\frac{\tau+1}{N}}} = O(\varepsilon^{\frac{6b}{N}}),$$

where

$$d(\varepsilon) = c(\varepsilon) \varepsilon^b = 2((2 + 3 + \cdots + \frac{1}{n-1}) \varepsilon^b + \frac{1}{c}).$$

Recall that $G_0 = G$, $\Lambda_0 = G_*$. Now let $G_\varepsilon = \Lambda_*$. Then by (3.2) and the measure estimate contained in the Normal Form Theorem, we have

$$|G \setminus G_\varepsilon| \leq |G_0 \setminus G_*| + |\Lambda_0 \setminus \Lambda_*| = O(\varepsilon^{\frac{\delta}{N}}).$$

The proof of our main result is now complete.

Acknowledgement. This work benefits a great deal from our discussions and communications with Professors Ken Meyer and Jesús Palacián. We would like to thank them for motivating us with interesting applications arising in celestial mechanics.

REFERENCES

- [1] V. I. Arnold, Small denominators and problems of stability of motion in classical mechanics, *Usp. Math. Nauk.* **18** (6) (1963), 91-192.
- [2] V. I. Arnold, Dynamical systems III, Springer-Verlag, Berlin Heidelberg, 1988.
- [3] L. Biasco, L. Chierchia, and D. Treschev, Stability of nearly Integrable, degenerate Hamiltonian dystems with two degrees of freedom, *J. Nonlinear Sci.*, **16** (2006), 79-107.
- [4] L. Biasco, L. Chierchia, and E. Valdinoci, N -dimensional elliptic invariant tori for the planar $(N + 1)$ -body problem, *SIAM J. Math. Anal.* **37** (2006), 2560-2588.
- [5] A. D. Bruno, Analytic form of differential equations. *Transactions of the Moscow Math. Soc.*, **25** (1971), 131-288.
- [6] A. Celletti and L. Chierchia, KAM stability for a three-body problem of the solar system, *Z. Angew. Math. Phys.* **57** (2006), 33-41.
- [7] C.-Q. Cheng and Y.-S. Sun, Existence of KAM tori in degenerate Hamiltonian systems. *J. Differential Equations* **114** (1994), 288-335.
- [8] S.-N. Chow., Y. Li, and Y. Yi, Persistence of invariant tori on submanifolds in Hamiltonian systems, *J. Nonl. Sci.*, **12** (2002), 585-617.
- [9] R. H. Cushman, A survey of normalization techniques applied to perturbed Keplerian systems, Dynamics Reported: Expositions in dynamical systems (Jones et al Ed.), Springer, Berlin, Heidelberg, 1992.
- [10] J. Féjoz, Quasiperiodic motions in the planar three-body problem, *J. Differential Equations* **183** (2002), 303-341.
- [11] J. Féjoz, Proof of ‘Arnold’s theorem’ on the stability of a planetary system (following Herman), *Ergod. Th. & Dynam. Sys.* **24** (2004), 1521-1582.
- [12] Y. C. Han and Y. Li, Arnold’s theorem on properly degenerate systems with the Rüssmann nondegeneracy, *Science in China ser. A Mathematics*, **48**(2005), 1656-1669.
- [13] Y. C. Han, Y. Li, and Y. Yi, Degenerate lower dimensional tori in Hamiltonian systems, *J. Differential Equations* **227** (2006), 670-691.
- [14] M. Kummer, On the regularization of the Kepler problem, *Commun. Math. Phys.* **84** (1982), 133-152.
- [15] M. Inarrea, V. Lanchares, J. F. Palacián, A. I. Pascual, J. P. Salas, and P. Yanguas, Reduction of some perturbed Keplerian problems, *Chaos Solitons Fractals* **27** (2006), 527-536.
- [16] Y. Li and Y. Yi, A quasi-periodic Poincaré’s Theorem, *Math. Annalen*, **326** (2003), 649-690.
- [17] Y. Li and Y. Yi, On Poincaré-Treshchev tori in Hamiltonian systems, *Proc. Equadiff 2003*, Dumortier et al (Ed.), World Scientific, 2005, 136-151.
- [18] Li, Y. and Yi, Y., Persistence of invariant tori in generalized Hamiltonian systems, *Ergod. Th. & Dyn. Sys.*, **22** (2002), 1233-1261.

- [19] U. Locatelli and A. Giorgilli, Construction of Kolmogorov's normal form for a planetary system, *Regul. Chaotic Dyn.* **10** (2005), 153-171.
- [20] M. Mazzocco, KAM theorem for generic analytic perturbations of the Euler system, *Z. Angew. Math. Phys.* **48** (1997), 193-219.
- [21] K. R. Meyer, Periodic Solutions of the N-Body Problem, Springer-Verlag, Berlin, Heidelberg, 1999.
- [22] K. R. Meyer, Scaling Hamiltonian systems, *SIAM J. Math. Anal.* **15** (1984), 877-889.
- [23] K. R. Meyer and D. S. Schmidt, From the restricted to the full three-body problem, *Trans. Amer. Math. Soc.* **352** (2000), 2283-2299.
- [24] Rüssmann, H., Invariant tori in non-degenerate nearly integrable Hamiltonian systems, *Regul. Chaotic Dyn.*, **6**(2001), 119-204.
- [25] H. Rüssmann, Nondegeneracy in the perturbation theory of integrable dynamical systems, *Number theory and dynamical systems* (York, 1987), 5-18, London Math. Soc. Lecture Note Ser., 134, Cambridge Univ. Press, Cambridge, 1989, *Stochastics, algebra and analysis in classical and quantum dynamics* (Marseille, 1988), 211-223, Math. Appl., 59, Kluwer Acad. Publ., Dordrecht, 1990.
- [26] Sevryuk, M. B., KAM-stable Hamiltonian, *J. Dynam. Control. Syst.*, **1**(1995), 351-366.
- [27] Sevryuk, M. B., Partial preservation of frequencies in KAM theory, *Nonlinearity*, **19** (2006), 1099-1140.
- [28] B. Sommer, A KAM Theorem for the Spatial Lunar Problem, Ph.D thesis, 2003.
- [29] J. Xu and J. You, Corrigendum for the paper Invariant tori for nearly integrable Hamiltonian systems with degeneracy, *Math. Z.* **257** (2007), 939.
- [30] J. Xu, J. You, and, Q. Qiu, Invariant tori for nearly integrable Hamiltonian systems with degeneracy, *Math. Z.*, **226** (1997), 375-387.

Y. HAN: SCHOOL OF MATHEMATICS, JILIN UNIVERSITY, CHANGCHUN, 130012, AND SCHOOL OF MATHEMATICS, SHANDONG UNIVERSITY, JINAN, 250100, P. R. CHINA

E-mail address: hanyc@jlu.edu.cn

Y. LI: SCHOOL OF MATHEMATICS, JILIN UNIVERSITY, CHANGCHUN, 130012, P. R. CHINA

E-mail address: liyong@jlu.edu.cn

Y. YI: SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA, GA 30332, USA, AND SCHOOL OF MATHEMATICS, JILIN UNIVERSITY, CHANGCHUN, 130012, P. R. CHINA

E-mail address: yi@math.gatech.edu